Maurizio Vichi (Sapienza University of Rome)

## New Methodological Challenges for the Era of Big Data

Topic 2 - Learning more from what we already know

Keywords: Big Data, Segmentation, Rankings, Composite Indicators

## Introduction

Big Data frequently describe complex economic, social and demographic phenomena evolving over time and having a territorial diffusion. These phenomena show different statistical relations such as spatial, cross-sectional, time series correlations, which need to be analysed in order to correctly read and extract the relevant information. Three or high dimensional arrays (data (iper)-cubes), are used to rearrange the huge number of statistical units with a spatial location, variables, and times. Phenomena described by this abundance of data need appropriate statistical tools to synthesize the relevant knowledge. A statistical modelling approach that takes into account the high dimensionality of the data, for learning more from what it is known from univariate analyses is used for segmenting the phenomena, in order to develop rankings, composite indicators and other tools useful for decision making.

## Methods / Problem statement

In this presentation, in particular, we illustrate methodologies for simultaneous segmentation including ranking of segments and hierarchical dimensional reduction of Big Data to obtain composite indicators. Methodologies are illustrate with examples. A final discussion follows.